
National Standards and Assessments: Will They Improve Education?

Author(s): Linda Darling-Hammond

Source: *American Journal of Education*, Vol. 102, No. 4, Educational Reform through National Standards and Assessment (Aug., 1994), pp. 478-510

Published by: The University of Chicago Press

Stable URL: <http://www.jstor.org/stable/1085465>

Accessed: 27/02/2010 21:08

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Education*.

National Standards and Assessments: Will They Improve Education?

LINDA DARLING-HAMMOND

Teachers College, Columbia University

National standards and assessments have been promoted as a means for upgrading curriculum and student performance in schools. This article argues that they are unlikely to have that effect for several reasons: First, top-down specifications of content linked to tests cannot take into account the many pathways to learning that will be appropriate for different students in schools across the country. Second, school communities must undertake their own hard work on standard setting and consensus development if they are to become committed to and knowledgeable about change. And, finally, large inequalities in opportunities to learn are more responsible for learning gaps than a paucity of tests. Standards and tests have already proved themselves to be an ineffectual means for leveraging resource equalization. Inequalities in learning opportunities must be addressed head-on if they are ever to be successfully removed. The article argues that, instead of starting with content and performance standards, policies should aim to create a system in which improved teacher knowledge and equalized school capacity are the starting points for systemic change. In such a system, teachers and schools will have the knowledge, resources, and organizational supports to create appropriate curriculum and useful assessments for the students they serve.

The adoption of ambitious national education goals by Congress and the National Governors Association has triggered nationwide debate on curriculum, testing, and national standards. Similar debates are occurring in states where curriculum and assessment reforms are aiming to promote more challenging forms of learning for students. The idea of promoting reform through the setting of standards for student learning is widespread (Sykes and Plastrik 1993; Clune 1993). One reason for this is that there has been a major shift in the kind of learning the society desires and requires of young people: from passive, rote learning to active engagement in problem solving, invention,

American Journal of Education 102 (August 1994)

© 1994 by The University of Chicago. All rights reserved.

0195-6744/94/0204-0005\$01.00

and management of information, resources, and ideas; from “basic” (read inert, “fill-in-the-blanks,” compliance-oriented) skills to “higher order” thinking and performance skills; from high levels of education for a few to a challenging curriculum for all (Cohen et al. 1993; Darling-Hammond 1993).

While proponents of differing views on these issues concur on the need for all students to aim for common learning goals and to achieve at high levels, disagreements exist over how these objectives can be promoted, how—and at what levels of the educational system—standards should be defined and their attainment measured, and how learning goals should be pursued and evaluated. Are national standards really likely to be more thoughtful and appropriate for shaping local school change than state or locally developed standards? Those who assume hierarchical intelligence (that higher levels of government are “smarter” than lower levels) would say yes: national content and performance standards will be sufficiently superior to local efforts that they should become the basis for shaping or certifying national, state, and/or local assessments (U.S. House 1993; Smith and O’Day 1990). Others argue that the power of local engagement in democratic discourse for standard setting is a more promising means for energizing productive change (Sizer 1991, 1992).

Other questions also arise. Is a national consensus about curriculum content and performance standards achievable or wise? Is there one right way to think about and configure what students know and should be able to do in a fashion that could serve all communities and students well? Those who argue for national standards view coherent guidance at the upper reaches of government as a necessary starting point for local consistency in planning and management (O’Day and Smith 1993). Others argue that, given the nature of knowledge and understanding and the diversity of human experiences and paths to learning, it is not plausible to decide from places far from schools and classrooms exactly what, how, when, and in what way ideas ought to be taught and student understanding tested (Clune 1993; Darling-Hammond 1993; Sizer et al. 1992–93).

LINDA DARLING-HAMMOND is William F. Russell Professor of Education and codirector of the National Center for Restructuring Education, Schools, and Teaching at Teachers College, Columbia University. Her research focuses on issues of school restructuring, teaching quality, and educational equity.

In addition to debates over the locus and governance of standards and assessments, there are even more deep-seated questions regarding how to translate ideas about learning into supportive and successful school practices. Can the mere issuance of standards really propel improvements in schooling, or are there other structural issues to contend with—issues such as funding, teachers' knowledge and capacities, access to curriculum resources, and dysfunctional school structures? Should students be the ones to bear the brunt of sanctions and consequences based on their performance, or should others take responsibility for the quality of education they provide to students—including states that fund schools unequally and districts that often favor schools with affluent and vocal parents over those with less powerful constituencies? If schools suffer sanctions for lower test scores, can we expect them to want to serve educationally needy students rather than pushing or keeping them out whenever possible?

Thus, a fundamental issue is whether new standards and assessment systems could support better teaching and transform schooling for all students, especially those who have been traditionally underserved by the system, or whether they will merely reify existing inequities. The answer to this question depends on the extent to which systemic reforms provide supports for effective teaching practices, promote equity in the allocation of educational resources, and enable more widespread school restructuring. It also depends on whether new standards and assessments are to be used primarily as traditional tests have been used, that is, for sorting and selecting students into or out of more and less rich curriculum and learning opportunities—or whether they are used to inform teaching that can be more adaptive to student learning styles and talents, strengths and needs (Glaser 1990; Gardner 1991).

This article argues that content standards aligned with tests are the wrong starting point for systemic school change aimed at improving teaching and learning for all students, and that national standards and assessments are the wrong vehicle. There are three reasons for this. First, top-down specifications of content linked to tests cannot take into account the many pathways to learning that will be appropriate for different students in schools across the country. There is no one curriculum Truth with a capital T that is waiting to be discovered and codified into a single set of guidelines superior to all others for all contexts and learning circumstances. Second, national standards and tests are inappropriate vehicles for enhancing teaching and stimulating school change. Building capacity for change requires that school communities undertake their own hard work on standard setting and

consensus development and that teachers be supported in their efforts to understand students, communities, and learning in diverse contexts, not just to implement external curriculum and testing mandates, regardless of their appropriateness for the learners they serve. And, finally, content and performance standards are already proving themselves, once again, to be a weak, ineffectual means for leveraging resource equalization. Inequalities in learning opportunities must be addressed head-on if they are ever to be successfully removed.

The constructivist learning theory that many content-standards advocates would like reflected in national standards creates a paradox for them: if students construct knowledge in highly contextualized ways based on their diverse, culturally grounded experiences, teaching must be highly adaptive, and curriculum must allow for many starting points and pathways. National content and performance standards aligned with grade-level tests (particularly if these are to be standardized and reliable enough to allow for calibration and certification judgments) will of necessity create a static and bounded conception of curriculum that is at odds with this understanding of learning. The idea that a top-down curriculum and testing system should be the primary guide for teachers' actions also betrays our understanding that teachers create curriculum based on a deep understanding of students and learning contexts as well as of subjects and curriculum goals (Darling-Hammond 1993).

At the point where curriculum standards are translated into tests, someone must decide who should have learned what in what fashion by what point in time. The assumption when this is done at a national level is that there is one particular configuration of useful knowledge and means for representing it that counts as Truth—one that is consensual and universally (or at least nationally) appropriate. This configuration then demands teaching that is aimed at straightforward implementation rather than at construction of knowledge—for students, for teachers, or for the profession and disciplines as a whole. In this view, as content standards are “adopted” or “certified” they become legislative tools. These are content standards that are not meant to be improved on with advances in knowledge or adapted as circumstances demand, they are to be implemented as top-down directives always are—more or less unthinkingly.

Rather than starting with static, politically negotiated specifications of curriculum content to guide state and local curriculum development, text adoptions, test development, teacher training, and other aspects of schooling, federal and state policies should aim to create a system in which improved teacher knowledge and school capacity are the starting points for systemic change. As testing expert George Ma-

daus asserts: “[T]eachers, not assessments, must be the cornerstone of any systemic reform directed at improving our schools. . . . [Policymakers] have lost sight of the fact that . . . “the teacher is a mediator between the knower and the known, between the learner and the subject to be learned. A teacher, not some [test or performance assessment], is the living link in the epistemological chain” (Palmer 1983, pp. 29–30; cited in Madaus 1993, p. 5).

Instead of specifying content and creating a system of technologies to enforce its transmission (Madaus 1993), the federal and, to a large extent, state governments’ role should be to create policies that ensure that teachers and schools have the knowledge, the material resources, and the support for school-based inquiry that will allow them to create good curriculum for the students they serve. Knowledgeable teachers working in schools that function as learning organizations would then be able to develop, select among, expand, and improve on any content or curriculum ideas offered from a variety of sources (including national professional standards), thus ensuring that knowledge advances are encouraged and incorporated, and that skillful teaching is adapted to individual students and school contexts.

As Wise (1979) argues, while equity concerns must be resolved by higher units of governance and must be centrally administered, productivity concerns, by their nature, cannot be effectively administered in a hierarchical fashion. Teaching knowledge must be used in highly individualized ways, while policies are necessarily uniform and standardized; thus, policy decisions about teaching and schooling processes cannot ever meet the demands of varying school and student circumstances (Shulman 1983; Darling-Hammond 1990a). Consequently, the proper “top-down” role is the provision of equitable and adequate learning opportunities to children by supporting the preparation and distribution of highly skilled teachers to all children and the provision of adequate schooling resources to support the school organizations in which these teachers will work with students. The proper “bottom-up” role, given these resources of knowledgeable teachers and adequate materials, is the development of contextually useful advances in curriculum and teaching quality (Wise 1979; Clune 1993; Darling-Hammond 1993).

If the goal of standard setting is the improvement of education for all children, rather than merely a more efficient means of sorting students and schools into “worthy” and “unworthy” categories, attention must be paid to building the capacity for schools to teach in the manner envisioned by these learning goals. This requires carefully developed policy efforts in the areas of teacher development, school development, and equalization of resources.

The Debate over National Content and Performance Standards

The discussion about national standards revolves around both content and performance standards, which are generally described as outlining what students should know and how well they should know it. In discussions during the last years of the Bush administration, the idea of a national test was proffered as a means of leveraging change and allocating some federal funds. That idea has been more or less replaced in current discussions by the idea of a system of examinations, developed at the regional, state, or local levels and calibrated or certified nationally. The *Goals 2000 Educate America* bill recently enacted by Congress seeks to create a body that will certify national, state, and local content and performance standards as well as assessments intended to measure achievement of the standards.

Proponents of using national standards and assessments as a mechanism for leveraging systemic change (O'Day and Smith 1993) believe that such standards are necessary to raise student performance across the nation. For example, as the National Council on Education Standards and Testing (NCEST) argued in its 1992 report:

In the absence of demanding content and performance standards, the United States has gravitated toward having a de facto minimal skills curriculum. The many state minimum competency tests, the lower level skills orientation of most textbooks, and state and local policies that do not adequately promote quality are examples of this minimal approach. . . . Such low expectations shortchange students and ill-serve the country. Yet as long as today's low standards remain in place, the performance of the majority of students is unlikely to improve substantially. [p. 12]

The Council's intent in recommending the establishment of national standards is to raise the ceiling for students who are currently above average and to lift the floor for those who now experience the least success in school, including those with special needs. States should work toward reducing gaps in students' opportunities to learn and in their performance, such as those now associated with race, income, gender, and geographical location. [P. 4]

This statement helps to illuminate the reasoning of many advocates of national standards and testing. It assumes that national standards will be higher or more rigorous than those developed at local and state levels, and that it is the absence of such standards that is a major cause of inadequate student performance. Furthermore, the statement expresses the view that, without higher standards, student performance will not improve, and that the presence of such standards will

create improved performance for both currently high-achieving and low-achieving students. The actual development of national curriculum standards and the effects of previous testing initiatives bring both of these presumptions into question.

O'Day and Smith (1993, p. 250) describe their proposal for "content-driven systematic school reform" as starting with "curriculum frameworks that establish what students should know and be able to do," followed by "alignment of education policies": teacher training directly aimed at learning to teach the specified content, curriculum materials and assessments tied to the frameworks, and increased flexibility for governance at the school level. The vision is that, "when fully implemented, this model of content-driven systemic reform would be a uniquely American adaptation of the educational policies and structures of many of the world's highly developed nations. It would marry the vision and guidance provided by coherent, integrated, centralized education policies common in many nations with the high degree of local responsibility and control demanded by U.S. tradition" (O'Day and Smith 1993, p. 252).

Given the messy state of educational governance and politics in this nation, the relative lack of professional vehicles for guiding school decision making, the conflicting policies to which school practitioners must respond (Darling-Hammond, 1990a), and the commercialization and balkanization of the enterprise, which allow commercially developed texts and tests to set standards, the argument for systemic change is an appealing one. AsSizer and colleagues note,

In approach and appeal, the partisans of this strategy seem the twenty-first century counterparts of the administrative progressives who transformed American schooling at the start of this century. . . . They too seem obsessed by the irrationality of the current system: its unreliable processes, the uneven quality of its products, its domination by parochial interests, and its unresponsiveness to current economy demands. . . . Their method proceeds by asking first what knowledge such productivity demands, and by then trying to map it back onto schooling's goals and practices. What remains problematic for them is not the finding and setting of standards, but the design of an efficient system capable of being driven by them. [Sizer et al. 1992–93, pp. 4–5]

This approach—alignment of school procedures against externally prescribed content specifications—is a continuation of the bureaucratic model of school reform that has predominated throughout the twentieth century, one that assumes that the hierarchical implementation of carefully specified procedures will yield the desired products

(students). “Based on faith in rationalistic organizational behavior, in the power of rules to direct human action, and in the ability of researchers to discover the common procedures that will produce desired outcomes, 20th century school reform has assumed that changing the design specifications for schoolwork will change the nature of education that is delivered in classrooms—and will do so in the ways desired by policy makers” (Darling-Hammond 1993, p. 754).

For many, the appeal of national standards and associated tests is that they provide the basis for new design specifications while also providing a lever for incentives that will grab the attention of schools. For example, many proponents argue that the national tests or systems of tests, once developed, can then be used to trigger rewards and sanctions to students, schools, and districts, thus “driving” reform. These proposals assume that schools can be made to improve by setting standards and creating incentives that will force school people to pay attention to them. An underlying assumption is that current educational problems exist largely because educators either don’t have precise enough targets to aim for, or because they—and their students—aren’t trying hard enough, or both. Supplying concrete goals along with carrots and sticks is the presumed answer to underperformance.

Tucker (quoted in “The Roundtable: A New ‘Social Compact’ for Mastery in Education” 1992), for example, argues that more high-stakes tests are needed to motivate students. It is the absence of such tests that has “systematically deprived our teachers of motivated students” (p. S3). Tucker and other advocates of test-driven reform (see, e.g., Hornbeck 1992) also argue for school-level sanctions to motivate schools to improve. The theory behind their proposals ignores the evidence that behaviorist models relying on extrinsic rewards and punishments “are not only increasingly discredited as valid and useful explanations for individual learning and actions, they have little hope of being transferable to the level of complex institutions” (Oakes 1993, p. 12; Sarason 1982; Senge 1992).

The theory that lack of test-based incentives is the cause of the nation’s educational woes also ignores the fact that over the last decade states and districts have enacted hundreds of high-stakes tests that already affect students’ chances for graduation, continued education, course placements, promotions, and a variety of life options. Substantial evidence indicates that these tests have not had the effect of “motivating” students or schools into greater learning (Darling-Hammond 1991; Jaeger 1991; Madaus 1992; McLaughlin 1991; Shepard 1991). In New York State, which is probably the most thoroughly tested state in the nation, with high stakes attached to tests scores from elementary school through graduation, the educationally counterproductive con-

sequences of high-stakes tests for learning, for teaching quality, and for access to challenging curriculum have been documented in an increasing number of studies (Allington and McGill-Franzen 1992; Smith 1986; Schoenfeld 1987; Price and Schwabacher 1992; Snyder et al. 1993). If high-stakes tests were all that we needed to create better education, we would have it by now, given the massive high-stakes testing industry this country already sustains (National Commission on Testing and Public Policy 1990).

Opponents of national standards and national testing raise a number of other concerns and issues: (1) that national standards will override local standard setting, which is a necessary aspect of democratic education required for stimulating a discourse about educational aims and for creating commitment to educational change and ownership of community ideals (Sizer 1991, 1992; Darling-Hammond and Ancess, in press); (2) that national standards and tests will necessarily narrow curriculum just as state and local tests already have in the United States, and as examinations have in other countries (Madaus 1988, 1993); and (3) that national tests will demonstrate, once again, how much disparity in educational achievement exists, but will do nothing to ameliorate unequal learning opportunities and may exacerbate unequal opportunities in education and employment by serving as additional arbiters of educational options, graduation, and employment.

Evidence demonstrates that curriculum and testing standards will not, in themselves, produce either greater quality or greater equity in opportunities to learn (Darling-Hammond 1991; Wise 1979). Greater access to more challenging and authentic learning opportunities are the real issue. These require more than tests. Equity in access to school resources must be tackled directly, along with investment in teachers' knowledge and their capacity for enlightened teaching. One critique of the NCEST report notes that the report "puts its faith in tests, maintaining that tests and standards 'can be the cornerstone of the fundamental, systemic reform necessary to reform schools.' Moreover, its specific proposals for national action are largely limited to tests and standards. Most of the remaining and more difficult aspects of school improvement, such as professional development and family and community supports are left for states and localities, and the report offers no specific proposals for dealing with them" (Koretz et al. n.d., pp. 1-2).

Increasingly, policymakers recognize that change requires not only motivation but also greater know-how and capacity for new forms of teaching and schooling. Equalization of resources and access to knowledge are both critical for building school capacity to teach stu-

dents in the ways that student performance standards require. This recognition has in turn triggered a debate on opportunity-to-learn (OTL) standards, which are proposed in the NCEST report under the term “delivery standards.”

Proponents of OTL standards argue that, since all schools do not now have the capacity to provide all students with the opportunities to learn that are necessary for achievement of the “old” standards, much less the “new” more demanding ones, a moral issue is raised. Critics of national standards and tests argue that, without redressing current inequalities, the current rhetoric about “world-class standards” and new kinds of assessment will perpetrate yet another cruel hoax on children in schools that haven’t the remotest chance of offering “world-class” education with the resources they command. If standard setting as a route to reform is to mean anything at all, it must mean supporting schools’ abilities to provide high-quality teaching and learning opportunities, which support students’ abilities to achieve the standards so many say they would like them to achieve.

Although the OTL debate has been raised, it has not succeeded in finding a firm foothold in the *Goals 2000* legislation, where national certification of standards and tests provides teeth for one side of the equation while general exhortations for state and local development of OTL standards provide almost none for the other side—the side that would support children in their learning. A recent pronouncement on OTL standards from the National Governors Association is equally weak. Whatever these standards might include, for the governors, equalization of dollars and resources is not prominent among them. Many proponents of national standards and aligned tests claim that, once these are in place, resource equalization and capacity building will follow. Critics point to a historical record that offers little optimism on this score, viewing the standards alignment strategy as a repeat of top-down curriculum reform efforts of the past, which tried to send ideas via curriculum packages and tests into schools that did not have the conceptual or material capacities to put them into effect.

Why Not Use National Standards and Tests to Drive Reform?

In an ideal world, curriculum standards developed by national professional associations would be helpful if they were used not as legislated guidelines, but as tools for informing (rather than directing) state and local curriculum building, teaching practice, and assessment development. Where they are well conceived and well constructed, they may

help local educators reflect on and evaluate their own efforts at shaping classroom- and school-level practices, as they incorporate contemporary knowledge about teaching and learning into statements of curriculum goals and discussions of teaching strategies. Particularly when conceptions of knowledge are changing, this kind of consensual effort among practitioners in a field can be useful to teachers as they rethink their curriculum, teaching, and assessment.

However, national curriculum standards are much less likely to be useful, and are much more likely to be harmful to teaching, learning, and school reform, if they are coupled to national assessments or are used as “test specifications” for state and local assessments. Used in this way, they are likely to create a national curriculum that will impede local reforms currently underway, stifle promising innovations, disconnect with the real learning needs of many students, suffer misuse at the hands of policymakers, and exacerbate inequalities in children’s access to learning opportunities.

At the point where national standards become linked to national assessments, they would cease to be a tool for stimulating local curriculum work and innovative thinking about school changes and become, in effect, a national curriculum. Rather than defining attributes of a thoughtful program, they begin to define the who, what, when, where, and how of learning, since formulating assessments requires drawing inferences about exactly what things students will learn and when they will learn these things in order for tasks or items to be constructed. This is much more problematic, as it assumes common contexts and conditions for learning that do not exist and standardizes practice when it needs to be flexible and grounded to be effective.

The Need to Acknowledge Contexts for Learning

In terms of improving education, curriculum standards at their best may provide a general aim and vision for changes that, if compatible with local aims, must still be worked through in more specific terms in the schools. This “working through” is in itself educative and supportive of change as it stimulates inquiry into local practices in relation to new standards that have engaged local ownership. It also allows the construction of curriculum that is connected to students’ experiences, cultures, and communities by building on prior knowledge and understandings in the course of reaching for new concepts and skills.

The constructivist learning theory that undergirds much of the rhetoric of the new standards work is itself the major argument for not nationalizing standards and assessments in a manner that would ulti-

mately prescribe a national curriculum. A curriculum that enables all students to learn must allow for differing starting points and pathways to learning so that students are not left out or left behind; allow for different strategies and approaches that meet varying learning styles and needs; allow for the reality that different areas of study are differentially relevant (and will be differently pursued) in various communities because of geographic, economic, topological, and cultural considerations; and allow for the prospect that students' demonstrations of their knowledge will be grounded in these contextual differences that vary widely in a country as large and as diverse as the United States.

Proponents of national testing in the United States often allege that national tests in other countries are a source of their educational successes. Yet most other countries with whom we compare ourselves do not have national tests or a national curriculum. Germany, Canada, and Australia, for example, use state or provincial examination systems where they have them at all. In the former West Germany, there were 11 state ministries of education that were setting separate *Abiturs* for each state. In France, there are 23 separate academies that set the *brevet de college* examinations and the *baccalauréat* exams for each academy (Koretz et al. n.d., p. 7). Schools in Great Britain have traditionally had the opportunity to select from among several regional assessment programs—even given the country's smaller size. Japan has no national governmental test, although there are college entrance examinations.

Even though most of these countries are the size of only one or two states in the United States, they still have enough diversity in views about education and contexts within which people live that curriculum and assessment also vary. In addition to the diverse examination systems that exist, schools and students may often choose which two or three examinations they will take out of an array of 30 or more that are offered (Madaus 1993). There are important social, political, and educational reasons for this diversity.

Sizer et al. make the argument for strong local involvement in standard setting in this way:

Finding and setting standards for American schools and American schoolchildren's performance is a highly problematic undertaking, inescapably so. Plato can't deliver the kind of standards needed, we think, nor can Adam Smith. This is because the values that must ground these standards are as diverse as the country, but they are also its unifying characteristics. These values are frequently ambiguous (as, for example, in the very case of whether we prefer individual freedom or civic virtue), changeable, and profoundly dependent on context. . . . They may indeed inhere

in certain visions of the ideal twenty-first century workplace or in the best practices of the disciplines, and indeed a pursuit of them there is a worthy undertaking. This does not mean, however, that they can be readily abstracted from these settings—the result of a two-year project, say, undertaken by scholars, teachers, and a federal agency or foundation. . . . To know where and how to seek standards, we think one must come to terms with the fact that good standards are not things which are clear, discrete, and fit for checklists. Much in the fashion of Aristotle, who claimed that essence is necessarily intertwined with experience, we believe that standards cannot exist apart from experience. To answer the question, “What is good enough here?” one must refer to images of *good enough*—the way people look, talk, act, or feel while being *good enough* in whatever performance they attempt. And in the process, one should not stray too far from where *here* is. [1992–93, pp. 6–7]

While it is possible to reach a knowledge-based consensus among members of the teaching profession about how children learn well and what the implications of that are for practice, questions of precisely what pieces of content children should learn, how, and when are in some sense irresolvable. Knowledge is exploding at an ever more rapid rate, and decisions about particular ways of construing that knowledge are always to some extent arbitrary. Fundamental concepts that structure a discipline are useful guideposts. Beyond those, however, “good” decisions about what content is important for students are grounded in the contexts in which people operate and to which they need to apply their knowledge. What is important knowledge for someone who lives in one region of the country or in a particular kind of community will be differentially important for somebody who lives somewhere else. Given that there is too much to be known, we will never derive a national answer that reveals the Truth about exactly which facts students should know or which subtopics they should study at particular moments in their lives.

When differences in living and learning experiences are ignored or wished away, they can also become sources of hidden bias in testing. One psychometrician testifying about cultural bias in testing gave as an example a set of reading questions from a test, developed in Minnesota but used in California, that happened to use sledding as a referent. Almost no children in southern California could answer the questions correctly as they had never seen or heard of a sled (S. Klein, personal communication). The example seems obvious but the problem is nearly unavoidable. All problems, items, and prompts have referents with which children from different communities will be differentially

familiar. Research illustrates that when test references differentially tap students' life experiences—whether these deal with regattas or piñatas—students' differential performances will be conditioned not by their underlying competence but by their contextual knowledge (Garcia and Pearson 1994).

In addition, useful and relevant areas of study will necessarily differ by context as well. Oceanography will be of greater interest to students and curriculum builders in coastal areas than to those in the Great Plains. Learning about farming and forest life may be of greater relevance to students in upper Wisconsin than the study of public transportation systems that engage children in New York City. Extended study of poisonous snakes may be important in the rural southwest and unimportant in downtown Chicago, where weather systems command greater curriculum time. And so on.

These variations are necessarily overlooked when curriculum and assessment decisions are made further away from the places students study. Tying national standards to national assessments will codify one version of knowledge and its applications in ways that could prevent teachers from creating the necessary bridges between students' diverse experiences and shared learning goals. This is especially problematic in the case of curriculum standards, such as those currently being developed in history, that are written in a prescriptive fashion, which chart content and teaching methods grade by grade as though all knowledge could be thus contained and as though all students would move in lockstep in attaining it. To be effective, teachers must meet students where they are, not where an idealized curriculum imagines they should be. Teachers must also be free to use material that is grounded in local referents that allow connections to be created. No national curriculum can ever do this.

Statements of curriculum goals would be more useful if they focused on how to prepare students to inquire successfully into new areas of study, to find and use information so that they will be able to analyze and generate ideas for themselves, to produce ideas and products so that they will have the tools, as new needs arise, to continually educate themselves for the world they will live in. In the long run this will be time much better spent than fighting about knowledge bytes as though there will be some final answer to the questions of whether, among the billions of pieces of information possible, students need to learn this one versus another one. The answers that satisfy people from district to district and state to state are likely to be different for credible and legitimate reasons, and this reality precludes the unbiased enforcement of a national curriculum and testing system.

The Difference between Professional and Governmental Efforts

Confidence in national standard setting as a strategy for reform was enhanced by the success of the National Council of Teachers of Mathematics' (NCTM) internally generated, profession-led activity to write curriculum standards. Begun over a decade ago, this effort differs in important ways from the standard-setting activities undertaken in other disciplines since. Its motives and intended uses, along with the process for development and extent of field involvement, model the possibilities for professional standards as tools for rethinking practice rather than as legislative sledgehammers for test-based reform.

The NCTM's long-range project was stimulated by a consensus among teachers that changes were needed in mathematics teaching. The standards for curriculum and for teaching have been introduced into schools and schools of education gradually and naturally, through professional networks and associations, as one stimulus among many for improving teaching. High school teacher Bill Fernekes explains the process this way, distinguishing it from some other recent standards-writing exercises: "NCTM not only garnered a tremendous amount of money to develop the standards, but they did it in a way that was intelligent, by saying they were going to take their time with the effort, develop the standards, and then work through their professional organization to spread the word. . . . [T]his hasn't happened through a top-down mandate. It has occurred through professional organizations and through publicity of those standards, done in a broad-based way" (Lockwood 1993, p. 12).

The standards have provided a useful and much-used set of ideas against which curriculum and teaching ideas are being tested and tried by local educators in a variety of ways across the country, supported by teacher-to-teacher networks like the Urban Mathematics Collaboratives and many others. The standards are organized around general constructs, learning goals, and areas of study, which provide a compass for change, but not a blueprint. They seek to get teachers pointed in the right direction but not to prescribe practice or to over-specify content. They differ from some of the more recent standards-writing activities in this regard: they illuminate the structure of the discipline and the goals of teaching and learning, but they do not prescribe when and how students must learn specific pieces of content.

Companion standards for teaching and teacher development point to NCTM's awareness that it is what teachers know and understand that is most critical to the future of mathematics teaching, not what any particular edicts for content standards might contain. Finally, the

dynamic nature of these standards is reflected in the fact that they are currently already under revision. Good standards always advance knowledge and, therefore, evolve of their own accord. Thus, they should not be legislated and thereby lodged in concrete.

The positive NCTM experience, along with a general predilection for top-down change, encouraged a faith in the rectitude of national standards that has spawned a set of federally funded efforts quite different in their genesis and approach. Most of the more recent standards development efforts were launched governmentally, rather than from a felt need for or consensus about change among teachers. Organizations given federal funds under the Bush administration—some of them representative of classroom teachers, others dominated by academics and noneducators—were charged with developing standards in other fields: geography, history, science, English, and, later, social studies, civics, the arts, and others. By some counts, there are now at least 13 national standards-writing activities underway.

On small budgets in most cases and on tight time frames in all cases, these groups have tried to hammer out standards like NCTM's without nearly the time or involvement from members in the field. These efforts vary greatly in the quality of the work, in the extent to which they incorporate a conception of curriculum that reflects current understandings about learning, in the extent to which they value deep understanding and complex performance versus rote recall of discrete factual material, and in the extent to which they represent a consensus in the field among scholars and practitioners.

Some of the new standards have already given rise to controversy. For example, the well-funded national history standards, modeled after a California curriculum framework that some districts in that state refused to use for its perceived insensitivity to multicultural contributions, is also governed by a content coverage orientation and a unidisciplinary approach that critics argue is at odds with the kind of in-depth study needed for students to engage in useful inquiry and to acquire real understanding. These standards are not easily compatible with the (much less well-funded) standards being developed by the National Council for the Social Studies or with those being developed by other organizations in geography and civics. Although they are offered in the name of systemic change aimed at a more coherent system, the standards are not written from a common conception of teaching, learning, or educational goals, much less of the disciplines themselves.

The emerging cacophony of standards, lacking the very coherence national standards advocates have wanted to promote, is a product of what I call the "assumption of hierarchical intelligence": that higher

levels of government makes superior decisions and create better-informed ideas than lower levels of government, and that these are applicable to and desirable for decisions at local levels.

The Flawed Assumption of Hierarchical Intelligence

This assumption has undergirded the mandate-oriented management of many federal social welfare programs and has created much of the state-level regulatory gridlock schools now experience. It is an assumption that has been increasingly brought into question not only by governmental alternatives that seek to stimulate local problem solving but also by the lessons of restructuring industries and businesses that have learned the importance of encouraging frontline inventiveness.

Because the NCTM effort (although professionally led and not federally stimulated or funded) has been helpful to the field of practice, it has strengthened the presumption in the Goals 2000 legislation that standards set nationally—even federally—will be somehow better or smarter than initiatives undertaken locally or at the state level, and, thus, should be the benchmark for “certifying” others. This assumption competes with ideals about local control and notions about decentralization, diversity, and invention. Experience shows that neither vision can be romantically embraced as always being applicable to the complexities of American life. But evidence thus far about the development of the recent batch of national standards under federal leadership belies the assumption of hierarchical intelligence and calls into question the desirability of requiring that localities evaluate their work against national benchmarks.

There are other problems with considering the national standards, once certified, as the basis for approving state and local standards and assessments. Many states and localities are already deeply engaged in developing their own curriculum standards and new assessments. Many of these initiatives are much further along than the national standards are, and many are based on more serious consensus building, better-grounded views of teaching and learning, and more sophisticated thinking about implementation and assessment. They illustrate that an effort undertaken federally does not ensure that it will be managed as well as or better than local initiatives. Unfortunately, local efforts may be held hostage to less thoughtful conceptions of disciplines, teaching, and schooling if they must be certified against national standards, rather than evaluated on their own terms.

While there are common threads and intentions across many of these efforts—such as an effort to encourage students to think critically and creatively rather than to parrot back material learned by rote—there are many diverse ways of achieving these goals. Schools are also exploring a variety of ways that they can reconfigure teaching and learning opportunities to provide for more intellectually challenging experiences for more students while also trying to provide that curriculum to a greater array of students and to personalize school structures. At this point of great experimentation, if national content and performance standards are tied to incentives and assessments, they may actually constrain those initiatives rather than support them.

As just one example, the new national standards are being written within traditional disciplines and even subdisciplines, which perpetuates older conceptions of how knowledge should be segmented and compartmentalized. They do not reflect the interdisciplinary perspective on teaching and learning that many state and local reforms are built on. Assessment reforms such as the New Standards Project and grassroots restructuring initiatives such as the Coalition of Essential Schools and the Foxfire Teacher Outreach Network are transforming school curriculum through interdisciplinary courses and assessments. Some states are developing curriculum frameworks in a more integrated fashion, including, for example, New York State's curriculum frameworks in "mathematics, science, and technology" and in "arts and humanities." With national standards determining the certification of local standards, would these be "uncertifiable" because they construe knowledge in a more interdisciplinary fashion than the federally sponsored projects?

The national standards also pose quandaries for those who envision these individual sets of standards as the complementary parts that would make up a whole curriculum. With 13 or more distinct sets of standards, teaching to all of them would require a "300 percent child" attending school several times longer and concentrating on no subject in very much depth. At a time when some schools are working to integrate learning, to press for in-depth understanding, to appreciate that "less is more," a mandate to incorporate all of these standards would be counterproductive.

While coherence is an important goal, it is most important that it be achieved at the local school level—where students and teachers must make sense of the educational environment, so that it works to enhance integrated and powerful learning. The only hope for this kind of local coherence is that educators be encouraged to use external stimuli such as national standards or state frameworks as tools for informing their thinking rather than as directives to be followed in

anticipation of accompanying high-stakes assessments. If schools are going to be encouraged to rethink how they do business by focusing on essential goals rather than on a cacophony of competing directives, state and federal policymakers will need to restrain themselves from thinking that changes must be mandated from the top, which would leave regulatory gridlock and intellectual chaos in the schools that must reconcile these conflicting impulses.

Building Capacity through Local Participation in Assessment

A final argument against linking national standards to national assessments has to do with the nature of local capacity building and the importance of involving teachers and other local educators in the processes of developing and implementing new assessments. For students to learn in different and more powerful ways, teachers must learn how to teach in different and more powerful ways. Teacher participation in assessment development is an important aspect of motivating and informing change efforts and of improving the quality of classroom teaching.

One aspect of this argument is that the locus and uses of testing must change if they are to be useful for supporting more informed teaching and more serious learning (Darling-Hammond 1994; Glaser 1990). Assessments must be developed and used in ways that help teachers understand students' learning and thinking so that they can support them more effectively. To the extent that standards and assessments continue to be used primarily as extrinsic arbiters of rewards and sanctions for students, rather than as means for providing enhanced learning opportunities for both students and teachers, they will do little to change the nature of education or of educational opportunity in this country.

Efforts to ensure that *all* students learn in meaningful ways that result in high levels of performance require that teachers know as much about students and their learning as they do about subject matter. However, teachers' understandings of students' strengths, needs, and approaches to learning are not well supported by external testing programs that send secret, secured tests into the school and whisk them out again for machine scoring that produces numerical quotients many months later.

Authentic assessment strategies can provide teachers with much more useful classroom information as they engage teachers in evaluating how and what students know and can do in real-life performance

situations. These kinds of assessment strategies create the possibility not only that teachers will develop curriculum aimed at challenging performance skills, but also that they will be able to use the resulting rich information about student learning and performance to shape their teaching in ways that can prove to be more effective for individual students (Darling-Hammond and Aneess, in press).

In addition, evidence from restructuring efforts indicates that, when schools wrestle with their own standard setting, the collective struggle to define directions, to evaluate progress, and to “map backward” to new curriculum and teaching possibilities can create an engine for schoolwide change that is absent when assessment is entirely externalized. If authentic forms of student assessment are shaped and implemented by members of the whole school community, they can enable the kinds of teacher, parent, and student learning that are needed to support the classroom and schoolwide changes required for student success (Darling-Hammond and Aneess, in press; Sizer et al. 1992–93; Darling-Hammond 1993).

The development of authentic assessment practices within schools can create a dynamic process of staff development and school development. This dynamic is set in motion as teachers struggle to articulate their goals and standards, to look for and create common ground in their views of subjects and students, to reflect on their practice, and to attend in new ways to learners and learning. The personal and organizational growth that occurs in this process depends on the intense engagement of staff in collaboratively defining, redefining, testing, and activating their own constructed and contextualized understanding of what is worth knowing and how it is worth assessing.

It is for this reason that the locus of assessment development and implementation is as important as the nature of the assessment tools and strategies. Assessments that are externally developed and scored cannot transform the knowledge and understanding of teachers and of school organizations—even if they are more performance based than are current tests. This is because teacher learning about the deeper structures of curriculum, about the nature and nuances of student thinking, and about the connections between teaching efforts and student performances derives substantially from firsthand, constructivist encounters with assessment development and from the subsequent evaluation of student work. Assessment reforms can increase student success by increasing organizational learning if they change not only the kinds of tasks students are asked to engage in but also the kinds of inquiry schools and teachers are called to undertake as they bring assessment into the heart of the teaching and learning process.

Developing Capacity for School Change and Improvement

What will it take for all schools to undertake seriously the development of a “thinking curriculum” for all students? I propose three key elements that ought to be the basis for a policy agenda at the national and state levels. The first is the development of a well-prepared teaching force that is made available to all students. This should be the starting point for systemic change, as it is the foundation on which productive learning opportunities must be built and through which a more challenging curriculum must be both developed and reinvented for particular classrooms and children.

The second is the assurance of resources and conditions for learning—the creation of equitable “opportunities to learn.” The third, closely related to the first two, is the creation of ongoing supports for school-based inquiry and change that create a press for continual improvement.

All of these areas of policy work intersect with standard-setting activities—with professional, voluntary curriculum standards developed from national, state, and local models; with performance standards adopted for various purposes by states, districts, and individual schools; and with standards for teacher knowledge, as well as state and local delivery systems. As Sykes and Plastrik (1993) describe, different kinds of standard setting can be used to advance change. In addition to curriculum standards and assessments for students, emerging professional standards for teachers and for teaching can help inform and transform school practice rather than merely shape testing.

Among these are new standards for teacher licensing and certification and for teacher education accreditation that have been developed to propel changes in teacher preparation and ongoing professional development (Darling-Hammond et al., in press). Developed through the National Board for Professional Teaching Standards, the Interstate New Teacher Assessment and Support Consortium, and the National Council for Accreditation of Teacher Education, these professional standards and assessments aim to ensure that teachers will have the skills, dispositions, and understandings needed to teach in the ways that new curriculum goals intend. In addition, professional organizations such as the National Association for the Education of Young Children and the National Council of Teachers of Mathematics have developed standards of practice to guide teaching and schooling.

These standards for teachers and teaching are linked to the curriculum standards being developed by many professional teaching associations by virtue of shared membership and explicit efforts to create a

unifying conception of teaching, learning, and professional responsibility. They are an important component of delivery standards that should guide policy-making regarding the parts of the system that must change for student learning to change—teacher education, licensing, certification, and staff development and school funding, organization, and reporting systems. It is in this way that conceptions of curriculum can inform the development of delivery standards, which include standards for teachers, standards for resources, and standards for practice that together provide students' opportunities to learn.

Inequality in Opportunities to Learn

In contrast to the more centralized and equalized supports for schools in other countries, there are still dramatic inequalities in education funding across schools in the United States, with wealthy schools commonly spending two or three times as much as poor ones (Educational Testing Service 1991; Kozol 1991). The resulting differences in educational experiences are most obvious in the contrasts between overcrowded, dilapidated schools with large classes in poor neighborhoods and bright, airy facilities where students study in well-equipped small classes in affluent neighborhoods.

Less visible but even more pronounced are inequalities in students' access to highly qualified teachers and high-quality curriculum. These inequalities start with unequal allocations of funds across and within states; they are compounded by unequal distributions of qualified teachers within districts and are further exacerbated by overt and tacit tracking within schools, which rations expert teachers and challenging curriculum to the more privileged students (College Board 1985; Oakes 1985, 1990; Darling-Hammond 1990*b*, in press).

Differences in school achievement can be traced to the effects of substantially different school opportunities offered to more and less advantaged students (Barr and Dreeben 1983; Dreeben and Gamoran 1986; Dreeben and Barr 1987; Oakes 1990). As Oakes (1990) found in looking at the distribution of mathematics and science opportunities across the nation: "Our evidence lends considerable support to the argument that low-income, minority, and inner-city students have fewer opportunities. . . . They have considerably less access to science and mathematics knowledge at school, fewer material resources, less-engaging learning activities in their classrooms, and less-qualified teachers. . . . Moreover, our findings are likely to be equally relevant for subject areas other than mathematics and science. The differences

we have observed are likely to reflect more general patterns of educational inequality" (pp. x–xi).

As a consequence of these inequalities, students are not only "at risk" from poverty, they are placed further at risk by the schools they attend. Because of teacher shortages, low-income students in many central city and poor rural schools are routinely taught by a disproportionate number of inexperienced and underprepared teachers, teachers teaching out of their fields of preparation, and a parade of short- and long-term substitutes hired when vacancies cannot be filled (Darling-Hammond 1990*b*, 1992). The differences in teacher qualifications across schools have startling effects on learning; they account more than any other factor for the differences in student achievement (Armour-Thomas et al. 1989; Ferguson 1991).

Because this nation has not invested heavily in teacher education and professional development, the capacity for the more complex approach to teaching envisioned by new standards is not prevalent throughout the current teaching force. Improving instruction for students requires improving preparation for *all* teachers, as well as ensuring an adequate supply of better-prepared teachers to all communities. To accomplish this, federal and state initiatives should seek to recruit and prepare new teachers, especially in shortage fields and in shortage locations, through scholarships and forgivable loans for high-quality teacher education; strengthen and improve teachers' preparation through supports for improving schools of education and through major licensing and certification reforms; and support teacher learning within schools by providing internships during the beginning teaching years, when 30–50 percent of new recruits drop out, and by developing learning networks for veteran teachers. If, in fact, the interaction between teachers and students is the most important aspect of effective teaching, then reducing inequality in learning has to rely on policies that provide equal access to competent, well-supported teachers.

Creating Opportunities to Learn

The report of NCEST, although devoted mainly to an argument for national performance standards for students, acknowledged the importance of qualified teachers and other resources to an accountability system that can actually stimulate school reform and improve education. The report called on states to select criteria for assessing schools' capacities and performances (National Council on Education Standards and Testing 1992, p. 13). This brief mention of standards to

promote improved education and equity was triggered by a thoughtful and persuasive report developed by the Council's Standards Task Force, which argued that school delivery standards are critical to the goals of a national standards system:

[I]f not accompanied by measures to ensure equal opportunity to learn, national *content* and *performance standards* could help widen the achievement gap between the advantaged and the disadvantaged in our society. If national *content* and *performance standards* and assessment are not accompanied by clear school delivery standards and policy measures designed to afford all students an equal opportunity to learn, the concerns about diminished equity could easily be realized. Standards and assessments must be accompanied by policies that provide access for all students to high quality resources, including appropriate instructional materials and well-prepared teachers. High content and performance standards can be used to challenge all students with the same expectations, but high expectations will only result in common high performance if all schools provide high quality instruction designed to meet the expectations. [National Council on Education Standards and Testing 1992, pp. E12–E13]

This concern was reinforced in the report of the NCEST Assessment Task Force, which listed indicators of school and system capacities as essential components of an assessment system, including evidence of “equal opportunity to learn curricula implied by the standards—learning resources, instruction, curriculum, teacher and student assignment” (National Council on Education Standards and Testing 1992, p. F13). The task force offered an important idea for linking the provision of educational opportunities to student performance standards, arguing that, as part of a comprehensive assessment system, states should collect evidence on system and school capacity paying specific attention to equity protection. Before states could use student performance assessments for any high-stakes purposes, such as student graduation or school interventions, they would have to provide evidence of both school capacity and assessment validity to a quality assurance agency. This agency would also conduct audits of schools to verify the capacity and equity evidence provided by states (National Council on Education Standards and Testing 1992, pp. F17–F18). This kind of strategy could leverage both school improvement and school equity reform, which would provide a basis for state legislation or litigation.

As Oakes (1989) argues, information about resources and school practices is essential “if [policymakers] want monitoring and account-

ability systems to mirror the condition of education accurately or to be useful for making improvements” (p. 182). Those who would attempt to use standards in the quest for accountability and improvement can themselves be held accountable for making sound decisions only if they address questions of *why* outcomes appear as they do and make necessary changes in the conditions that influence learning.

Opportunity-to-learn standards should point the attention of policy-makers and educators to those resources and learning experiences required for students to achieve the intellectually and practically challenging learning outcomes envisioned by current school reform initiatives. They should be used to create information and incentives for states, school districts, schools, teachers, and parents to act on in improving and equalizing students’ learning opportunities. The adoption and use of such standards should

1. provide information about the nature of the teaching and learning opportunities made available to students in different districts, schools, and intraschool programs across the state,
2. shape state legislation to create equity in funding and to distribute highly qualified teachers to local districts so that all students have access to an excellent education,
3. provide incentives to school districts to create policies that will ensure adequate and equitable resources, curriculum opportunities, and expert teaching to all of the schools within their jurisdiction,
4. help state education departments and local education agencies identify schools that need support in developing, progressing toward, and achieving adequate and equitable opportunities to learn for all of their students,
5. help schools and districts engage in self-assessment and peer review of practice in light of (a) the curriculum entitlements implied by state and locally defined learning outcomes and (b) the nature of teaching and learning activities reflected in professional standards of practice, and
6. provide local communities and parents with information on what constitutes adequate and equitable opportunity-to-learn standards so that they can use that knowledge in the interests of their children’s education and so that they can more effectively participate in the decision-making process of their local schools and districts.

A major challenge in developing standards that can be used to evaluate students’ opportunities and to trigger appropriate policy attention is how to configure expectations and actions so that they improve opportunities without imposing undue restraints on schools. The stan-

dards and their uses must be designed so that they can establish meaningful indicators and guidance for states, districts, and schools while being flexible enough to accommodate many different strategies for providing high-quality, appropriate education.

This tension may be resolved by envisioning two different kinds of standards. First, a small number of quantifiable resource measures may be evaluated through straightforward indicators. So long as these are at a sufficiently general level that they do not constrain decisions about how to configure resources to deliver education, they may avoid the problem of overspecification while still promoting equity. These *standards for delivery systems* should create incentives for states, school districts, and schools to ensure the equitable availability of such basic resources as funding, highly qualified teachers, and educational facilities and materials to all students.

Second, school and classroom practices are better understood by looking at them through an expert lens that can gauge the appropriateness of diverse approaches, rather than by requiring standardized measures and approaches. Education can learn from the ways in which school inspectorate systems operate in countries like England and peer review processes operate in professions such as medicine, accounting, and law, where standards of practice are part of accreditation and review processes. A school review that focuses on examining students' learning opportunities can evaluate standards of practice without resorting to inappropriate, standardized measures that might be gathered from afar in a traditional bureaucratic system of regulation.

Standards for practice should be based on professional knowledge about the teaching and schooling conditions under which children learn well. Examples of professional standards for schooling and curriculum include those promulgated by the National Association for the Education of Young Children and by the professional subject matter associations, such as the National Council of Teachers of Mathematics. These standards of practice, along with meaningful standards for preparation, licensure, and evaluation of practitioners, are a key component of professional accountability. They can be used to guide school-based processes of inquiry, self-evaluation, consultation, and problem solving, as well as external reviews by teams of practitioners.

In New York State, this approach to developing opportunity-to-learn standards is being implemented through three interrelated processes: (1) the articulation of *desired learning outcomes* at the state and local levels and the development of associated curriculum frameworks that articulate a curriculum entitlement for students as reflected in the learning outcomes; (2) the equalization of *school funding* and evaluation of school resource adequacy in light of this curriculum entitle-

ment; and (3) the development of a *school quality review* process that examines and supports school development through cycles of external and internal peer review focused on teaching and learning opportunities.

A framework for evaluating standards of practice has been articulated by Darling-Hammond (1992) and has been elaborated through the New York School Quality Review Initiative currently being piloted in over 50 schools. A similar practitioner-led review process is also under development in California and is under consideration in other states. The “lenses” for examining practice in the New York school quality review include the following:

1. School climate: students should encounter an environment that is respectful, purposeful, physically and psychologically safe, and personalized so as to ensure close, sustained relationships between students and teachers and attention to special needs.
2. Curriculum and assessment: all students should have access to a rich and challenging curriculum that fosters their critical thinking and creative and performance capabilities, that develops their multiple intelligences and diverse talents, and that encourages them to apply their learning in problem-solving situations. This should be supported by assessment strategies that are appropriate and authentic measures of the goals being pursued as well as instructionally useful indicators of individual student growth.
3. Teaching and learning experiences: students should be taught in ways that are cognitively and developmentally appropriate, and that respect their individual experiences, learning styles, and learning needs. Teachers should use a wide array of teaching strategies and adapt their teaching to support student success. Student access to curricular opportunities should be inclusive and adaptive for maximum talent development rather than exclusive for purposes of limiting enrollment or participation.
4. Professional inquiry and development: schools should have mechanisms that help them to continually evaluate how well they are meeting students’ needs and to provide ongoing supports for parent involvement, as well as for teacher development, consultation, and school improvement.

In the school quality review process, teams of expert practitioners look intently at teaching and learning experiences over a weeklong visit, reflect back to school members what they see, and help to stimulate thinking about ongoing school development and improvement. Similarly, a number of grassroots reform organizations, such as the Coalition of Essential Schools, the School Development Program, and the Accelerated Schools Program have organized their work around

common principles; where these are made tangible through school review or “critical friend” processes, they function like these standards of practice. Along with opportunity-to-learn standards that hold states, districts, and schools accountable for providing resources, these kinds of efforts can stimulate school growth and inquiry.

Establishing a commitment to collective problem solving that will permeate the school requires ongoing inquiry in light of locally valued goals. This commitment is undermined when all of the important evaluation and assessment functions are externalized. An inquiry ethos must be supported by methods for continually evaluating what is going on, asking not just *what* is occurring but also *why* it is happening and *whether* existing practices are accomplishing what the school community wants to accomplish. These questions should not be raised only once or twice a year, when the students are tested or the annual needs assessment form is filled out. They should be raised in every faculty and team meeting, on every occasion when faculty and students are striving to meet their goals, and at every juncture at which any kind of stocktaking occurs.

In some rare schools, this kind of collective questioning and reflection is frequent. In others, which are more common, it is rare. Teacher isolation has worked against collective accountability, while centralized planning, decision making, and evaluation have often removed occasions as well as incentives for this kind of activity at the school level. Yet, if schools are to become more responsive and open to change, they must find ways—as other professional organizations do—to make evaluation and assessment part of their everyday lives. Just as hospitals have standing committees of staff that meet regularly to discuss the effectiveness of various aspects of the hospital’s functioning, so schools must have regularized occasions for examining their practices. Creating ongoing occasions for self-assessment and evaluation through peer and community review of practice is a critical aspect of school improvement.

The systemic changes sought by investments in teacher knowledge, school resources, and school inquiry must also connect with efforts to define what students should know and be able to do, and with initiatives to rigorously assess student learning. In states such as New York, Vermont, Connecticut, and California, carefully targeted state assessments at key developmental points will provide data for informing policymakers about program successes and needs, about areas where assistance and investment are needed, and about assessment models for local schools. Meanwhile, locally implemented assessment systems—including portfolios, projects, performance tasks, and structured teacher observations of learning—will provide the multiple

forms of evidence about student learning needed to make sound judgments about instruction. In these models, assessment is used as a learning tool for schools and teachers rather than as an external lever for change. Standards and assessments are developed where people can look each other in the eye and argue about what's good enough and how they can collectively know that it's occurring.

Many schools, districts, reform networks, and professional organizations have already made inroads in the development of such assessments and in their use for supporting teaching and learning. Strategies for assessing learning through exhibitions, portfolios, projects, and careful observations of children have been invented and shared among grassroots school reform initiatives stimulated by such organizations as the Coalition of Essential Schools, Project Zero, the Foxfire Teacher Outreach Network, the North Dakota Study Group, the Prospect Center, and other networks of progressive schools, along with organizations such as the National Association for Education of Young Children, the National Council of Teachers of Mathematics, and other professional associations.

These approaches to assessment development have aimed at strengthening teaching and learning at the school level by engaging students in more meaningful, integrative, and challenging work and by helping teachers to look carefully at performance, to understand how students are learning and thinking, to reflect on student strengths and needs, and to support them with adaptive teaching strategies.

Where these efforts are underway, changes in teaching and schooling practices occur—especially for students who are not as often successful at schoolwork (Kornhaber and Gardner 1993; Darling-Hammond and Ancess, in press; Falk and Darling-Hammond 1993). Teachers' engagement in developing standards and assessments and in evaluating student work has supported their learning about student learning. As teachers learn about how students approach tasks, what helps them learn most effectively, and what assessment tasks challenge and support the kinds of learning desired, they find themselves transforming both their teaching and their assessment strategies. The more information teachers obtain about what students know and think as well as about how they learn, the more capacity they have to rethink their pedagogy, and the more opportunities they create for student success.

At the policy level, the top-down, content-driven approach to standards and the bottom-up approach driven by investments in teacher and school learning represent different theories of organizational change. One view seeks to induce change through prescriptions for practice buttressed by extrinsic rewards and sanctions for schools and

students, on the assumption that the fundamental problem is a lack of will to change on the part of educators. The other view assumes that the fundamental problem is a lack of knowledge about the possibilities for teaching and learning, combined with an undeveloped organizational capacity for ongoing change and improvement. This view seeks to induce change by building knowledge among school practitioners and parents about alternative methods, by stimulating organizational rethinking through opportunities to work together on the design of teaching and schooling, and by providing the means to try new approaches.

The developmental view of assessment seeks to create the conditions that enable responsible and responsive practice, which include teacher knowledge, school capacity for improvement and problem solving, flexibility in meeting the actual needs of real people, shared ethical commitments among staff, and appropriate policy structures that encourage rather than punish inclusive education (Darling-Hammond and Snyder 1992).

Engaging teachers in assessment is a critical aspect of that process. That engagement becomes a powerful vehicle for professional development, for supporting teachers in looking at and understanding student learning, for investigating the effects of teaching on learning, and for transforming their practices so that they become more effective. It is this insight into what students are really doing, thinking, and learning that is one of the greatest contributions of a locally grounded standards and assessment initiative to teacher development—and, hence, to student learning.

References

- Allington, Richard L., and Anne McGill-Franzen. "Unintended Effects of Educational Reform in New York." *Educational Policy* 6 (1992): 397–414.
- Armour-Thomas, E., C. Clay, R. Domanico, K. Bruno, and B. Allen. *An Outlier Study of Elementary and Middle Schools in New York City: Final Report*. New York: New York City Board of Education, 1989.
- Barr, Rebecca, and Robert Dreeben. *How Schools Work*. Chicago: University of Chicago Press, 1983.
- Clune, William H. "The Best Path to Systemic Educational Policy: Standard/Centralized or Differentiated/Decentralized?" *Educational Evaluation and Policy Analysis* 15 (1993): 233–54.
- Cohen, David K., Milbrey W. McLaughlin, and Joan E. Talbert, eds. *Teaching for Understanding: Challenges for Policy and Practice*. San Francisco: Jossey-Bass, 1993.
- College Board. *Equality and Excellence: The Educational Status of Black Americans*. New York: College Entrance Examination Board, 1985.

- Darling-Hammond, Linda. "Instructional Policy into Practice: 'The Power of the Bottom over the Top,'" *Educational Evaluation and Policy Analysis* 12 (Fall 1990): 233–42. (a)
- Darling-Hammond, Linda. "Teacher Quality and Equality." In *Access to Knowledge: An Agenda for Our Nation's Schools*, edited by John Goodlad and Pamela Keating. New York: College Entrance Examination Board, 1990. (b).
- Darling-Hammond, Linda. "The Implications of Testing Policy for Quality and Equality." *Phi Delta Kappan* 73 (1991): 220–25.
- Darling-Hammond, Linda. *Standards of Practice for Learner-centered Schools*. New York: National Center for Restructuring Education, Schools, and Teaching, Teachers College, Columbia University, 1992.
- Darling-Hammond, Linda. "Reframing the School Reform Agenda: Developing Capacity for School Transformation." *Phi Delta Kappan* 74 (June 1993): 753–61.
- Darling-Hammond, Linda. "Performance-based Assessment and Educational Equity." *Harvard Educational Review* 64 (Spring 1994): 5–30.
- Darling-Hammond, Linda. "Inequality and Access to Knowledge." In *Handbook of Multicultural Education*, edited by James A. Banks. New York: Macmillan, in press.
- Darling-Hammond, Linda, and Jacqueline Ancess. "Authentic Assessment and School Development." In 93rd yearbook of the National Society for the Study of Education, edited by Dennie Palmer Wolf and Joan Boyckoff Baron. Chicago: University of Chicago Press, in press.
- Darling-Hammond, Linda, and Jon Snyder. "Reframing Accountability for Learner-centered Practice." In *The Changing Contexts of Teaching*, edited by Ann Lieberman. 92nd yearbook of the National Society for the Study of Education, edited by Ann Lieberman. Chicago: University of Chicago Press, 1992.
- Darling-Hammond, Linda, Arthur E. Wise, and Stephen Klein. *A License to Teach: The Foundation for 21st Century Schools*. Boulder, Colo.: Westview, in press.
- Dreeben, Robert, and Rebecca Barr. "Class Composition and the Design of Instruction." Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., 1987.
- Dreeben, Robert, and Adam Gamoran. "Race, Instruction, and Learning." *American Sociological Review* 51 (1986): 660–69.
- Educational Testing Service. *The State of Inequality*. Princeton, N.J.: Educational Testing Service, 1991.
- Falk, Beverly, and Linda Darling-Hammond. *The Primary Language Record at P.S. 261*. New York: National Center for Restructuring Education, Schools, and Teaching, Teachers College, Columbia University, 1993.
- Ferguson, Ronald F. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation* 28 (Summer 1991): 465–98.
- Garcia, Georgia, and David Pearson. "Assessment and Diversity." In *Review of Research in Education*, vol. 20, edited by Linda Darling-Hammond. Washington, D.C.: American Educational Research Association, 1994.
- Gardner, Howard. *The Unschooled Mind*. New York: Basic, 1991.
- Glaser, Robert. *Testing and Assessment: O Tempora! O Mores!* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1990.
- Hornbeck, David. "The True Road to Equity." *Education Week* 11 (May 6, 1992): 32, 25.

- Jaeger, Richard M. "Legislative Perspectives on Statewide Testing." *Phi Delta Kappan* 73 (1991): 239–42.
- Koretz, Daniel M., George F. Madaus, Edward Haertel, and Albert E. Beaton. *National Educational Standards and Testing: A Response to the Recommendations of the National Council on Education Standards and Testing*. Congressional testimony. Santa Monica, Calif.: RAND, n.d.
- Kornhaber, Mindy, and Howard Gardner. *Varieties of Excellence: Identifying and Assessing Children's Talents*. New York: National Center for Restructuring Education, Schools, and Teaching, Teachers College, Columbia University, 1993.
- Kozol, J. *Savage Inequalities*. New York: Crown, 1991.
- Lockwood, Anne Turnbaugh. "National Standards: Who Benefits?" *Focus in Change* 11 (1993): 1–14.
- McLaughlin, Milbrey. "Test-based Accountability as a Reform Strategy." *Phi Delta Kappan* 73 (1991): 248–50.
- Madaus, George F. "The Influence of Testing on Curriculum." In *Critical Issues in Curriculum*, edited by Laurel Tanner. Chicago: University of Chicago Press, 1988.
- Madaus, George F. "The Effects of Important Tests on Students: Implications for a National Examination System." *Phi Delta Kappan* 73 (1991): 226–31.
- Madaus, George F. "A National Testing System: Manna from Above? A Historical/Technological Perspective." *Educational Assessment* 1 (1993): 9–26.
- National Commission on Testing and Public Policy. *Reforming Assessment: From Gatekeepers to Gateway to Education*. Chestnut Hill, Mass.: Boston College, 1990.
- National Council on Education Standards and Testing. *Raising Standards for American Education*. Washington, D.C.: Government Printing Office, 1992.
- Oakes, Jeannie. *Keeping Track*. New Haven, Conn.: Yale University Press, 1985.
- Oakes, Jeannie. "What Educational Indicators? The Case for Assessing the School Content." *Educational Evaluation and Policy Analysis* 11 (Summer 1989): 181–99.
- Oakes, Jeannie. *Multiplying Inequalities: The Effects of Race, Social Class, and Ability Grouping on Access to Science and Mathematics Education*. Santa Monica, Calif.: RAND, 1990.
- Oakes, Jeannie. *New Standards and Disadvantaged Schools*. Paper prepared for research forum, Effects of New Standards and Assessments on High Risk Students and Disadvantaged Schools, Harvard University, October 1993.
- O'Day, Jennifer A., and Marshall S. Smith. "Systemic School Reform and Educational Opportunity." In *Designing Coherent Education Policy: Improving the System*, edited by Susan Fuhrman. San Francisco: Jossey-Bass, 1993.
- Palmer, P. J. *To Know as We Are Known: A Spirituality of Education*. San Francisco: Harper & Row, 1983.
- Price, Janet, and Sara Schwabacher. *The Multiple Forms of Evidence Study*. New York: Fund for New York City Public Education, 1992.
- Sarason, Seymour. *The Culture of the School and the Problem of Change*. Boston: Allyn & Bacon, 1982.
- Schoenfeld, Alan H. "When Good Teaching Leads to Bad Results: The Disasters of 'Well Taught' Mathematics Courses." In *Educational Psychologist* ("Special Issue: Learning through Instruction: The Study of Students' Thinking during Instruction in Mathematics," edited by Penelope L. Peterson and Thomas L. Carpenter) (1987).

- Senge, Peter M. "Building Learning Organizations." *Journal for Quality and Participation* (March 1992). Reprint, Framingham, Mass.: Innovation.
- Shepard, Lorrie A. "Will National Tests Improve Student Learning?" *Phi Delta Kappan* 73 (1991): 232–38.
- Shulman, Lee. "Autonomy and Obligation." In *Handbook of Teaching and Policy*, edited by Lee Shulman and Gary Sykes. New York: Longman, 1983.
- Sizer, Theodore. "No Pain, No Gain." *Educational Leadership* (May 1991), pp. 32–34.
- Sizer, Theodore. *Horace' School: Redesigning the American High School*. Boston: Houghton Mifflin, 1992.
- Sizer, Theodore R., Joseph P. McDonald, and Bethany Rogers. "Standards and School Reform: Asking the Basic Questions." *Stanford Law and Policy Review* 4 (Winter 1992–93): 27–35.
- Smith, Frank. *High School Admission and the Improvement of Schooling*. New York: New York City Board of Education, 1986.
- Smith, Michael, and Jennifer O'Day. "Systemic School Reform." In *The Politics of Curriculum and Testing*, edited by Susan H. Fuhrman and Barbara Malen. 1990 yearbook of the Politics of Education Association. London: Taylor & Francis, 1990.
- Snyder, Jon, Edward Chittenden, and Priscilla Ellington. *Assessment of Children's Reading: A Comparison of Sources of Evidence*. New York: National Center for Restructuring Education, Schools, and Teaching, Teachers College, Columbia University, 1993.
- Sykes, Gary, and Peter Plastrik. *Standard-Setting as Educational Reform*. Washington, D.C.: ERIC Clearinghouse on Teacher Education, American Association of Colleges for Teacher Education, 1993.
- "The Roundtable: A New 'Social Compact' for Mastery in Education." Special report. *Education Week*, June 17, 1992.
- Wise, Arthur E. *Legislated Learning*. Berkeley and Los Angeles: University of California Press, 1979.